Computational Models of Historical Scientific Discoveries

Pat Langley, Institute for the Study of Learning and Expertise
Lorenzo Magnani, Department of Philosophy, University of Pavia
Peter C.-H. Cheng, School of Psychology, University of Nottingham
Adrian Gordon, Department of Computing, University of Northumbria
Sakir Kocabas, Space Engineering Department, Istanbul Technical University
Derek H. Sleeman, Department of Computing Science, University of Aberdeen

The discovery of scientific knowledge is one of the most challenging tasks that confront humans, yet cognitive science has made considerable progress toward explaining this activity in terms of familiar cognitive processes like heuristic search (e.g., Langley et al., 1987). A main research theme relies on selecting historical discoveries from some discipline, identifying data and knowledge available at the time, and implementing a computer program that models the processes that led to the scientists' insights. The literature on computational scientific discovery includes many examples of such studies, but initial work in this tradition had some significant drawbacks, which we address in this symposium.

One such limitation was that early research in law discovery ignored the influence of domain knowledge in guiding search. For example, Gordon et al. (1994) noted that attempts to fit data from solution chemistry in the late 1700s took into account informal qualitative models like polymerization and dissociation. They have developed HUME, a discovery system that draws on such qualitative knowledge to direct its search for numeric laws. HUME utilizes this knowledge not only to rediscover laws found early in the history of solution chemistry, but also to explain, at an abstract level, the origins of other relations that scientists proposed and later rejected.

Early discovery research also downplayed the role of diagrams, which occupy a central place in many aspects of science. For example, Huygens' and Wren's first presentations of momentum conservation took the form of diagrams, suggesting they may have been instrumental in the discovery process. In response, Cheng and Simon (1992) have developed HUYGENS, a computational model for inductive discovery of this law that uses a psychologically plausible diagrammatic approach. The system replicates the discovery by manipulating geometric diagrams that encode particle collisions and searching for patterns common to those diagrams. The quantitative data given to the system are equivalent to those available at the time of the original discovery.

Another challenge concerns the computational modeling of extended periods in the history of science, rather than isolated events. To this end, Kocabas and Langley (1995) have developed BR-4, an account of theory revision in particle physics that checks if the current theory is consistent (explains observed reactions) and complete (forbids unobserved reactions), revises quantum values and posits new particles to maintain consistency, and introduces new properties to maintain completeness. BR-4 models, in abstract terms, major developments in particle physics over two decades, including the proposal of baryon and lepton numbers, postulation of the neutrino, and prediction of numerous reactions. Background knowledge about symmetry and conservation combine with data to constrain the search for an improved theory in a manner consistent with the incremental nature of historical discovery.

We hope this symposium will encourage additional research that extends our ability to model historical scientific discoveries in computational terms.

References

Cheng, P. C.-H., & Simon, H. A. (1992). The right representation for discovery: Finding the conservation of momentum. *Proceedings of the Ninth International Conference on Machine Learning* (pp. 62–71). San Mateo, CA: Morgan Kaufmann.

Gordon, A., Edwards, P., Sleeman, D., & Kodratoff, Y. (1994). Scientific discovery in a space of structural models. *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society* (pp. 381–386). Atlanta: Lawrence Erlbaum.

Kocabas, S., & Langley, P. (1995). Integration of research tasks for modeling discoveries in particle physics. *Proceedings of the AAAI Spring Symposium on Systematic Methods of Scientific Discovery* (pp. 87–92). Stanford, CA: AAAI Press.

Langley, P., Simon, H. A., Bradshaw, G. L., & Żytkow, J. M. (1987). *Scientific discovery: Computational explorations of the creative processes*. Cambridge, MA: MIT Press.